Genomewide Association Studies: History, Rationale, and Prospects for Psychiatric Disorders

Cichon, S.; Craddock, N.; Daly, M.J.; Faraone, S.V.; Gejman, P.V.; Kelsoe, J.; Lehner, T.; Levinson, D.F.; Moran, A.P.; Sklar, P.; Sullivan, P.F.; Boomsma, D.I.; de Geus, E.J.C.; Posthuma, D.; Willemsen, G.

published in
American Journal of Psychiatry
2009

DOI (link to publisher)
10.1176/appi.ajp.2008.08091354

Genomewide Association Studies: History, Rationale, and Prospects for Psychiatric Disorders

Psychiatric GWAS Consortium Coordinating Committee

Objective: The authors conducted a review of the history and empirical basis of genomewide association studies (GWAS), the rationale for GWAS of psychiatric disorders, results to date, limitations, and plans for GWAS meta-analyses.

Method: A literature review was carried out, power and other issues discussed, and planned studies assessed.

Results: Most of the genomic DNA sequence differences between any two people are common (frequency >5%) single nucleotide polymorphisms (SNPs). Because of localized patterns of correlation (linkage disequilibrium), 500,000 to 1,000,000 of these SNPs can test the hypothesis that one or more common variants explain part of the genetic risk for a disease. GWAS technologies can also detect some of the copy number variants (deletions and duplications) in the genome. Systematic study of rare variants will require large-scale resequencing analyses. GWAS methods have detected a remarkable number of robust genetic associations for dozens of common diseases and traits, leading to new pathophysiological hypotheses, although only small proportions of genetic variance have been explained thus far and therapeutic applications will require substantial further effort. Study design issues, power, and limitations are discussed. For psychiatric disorders, there are initial significant findings for common SNPs and for rare copy number variants, and many other studies are in progress.

Conclusions: GWAS of large samples have detected associations of common SNPs and of rare copy number variants with psychiatric disorders. More findings are likely, since larger GWAS samples detect larger numbers of common susceptibility variants, with smaller effects. The Psychiatric GWAS Consortium is conducting GWAS meta-analyses for schizophrenia, bipolar disorder, major depressive disorder, autism, and attention deficit hyperactivity disorder. Based on results for other diseases, larger samples will be required. The contribution of GWAS will depend on the true genetic architecture of each disorder.

(Am J Psychiatry 2009; 166:540–556)

Since 2005 (1), genomewide association studies (GWAS [jē´ wōs]) have produced strongly significant evidence that specific common DNA sequence differences among people influence their genetic susceptibility to more than 40 different common diseases (2). Many of these findings implicate previously unsuspected candidate genes and new pathophysiological hypotheses. The method is feasible because millions of human DNA sequence variations have been catalogued and new technologies have been developed that can assay more than one million variants rapidly and accurately. The first GWAS reports have appeared for psychiatric disorders, and nearly 50 GWAS are completed or will be reported soon for attention deficit hyperactivity disorder (ADHD), autism, bipolar disorder, major depressive disorder, and schizophrenia. Additional studies are in progress. We formed an international consortium of psychiatric GWAS investigators to carry out rapid meta-analyses of these five disorders to maximize power. In the present overview, we describe GWAS methods, their rationale, and current results for nonpsychiatric and psychiatric disorders as well as limitations and uncertainties.

Candidate Genes, Linkage, and Linkage Disequilibrium

Genetic Epidemiology

Before any molecular genetic study is undertaken, the methods of genetic epidemiology are used to identify a phenotype (observable disease or trait) that is at least partially heritable. An introduction to these methods is available online (http://www.dorak.info/epi/genetepi.html). Briefly, twin, family, and population-based studies are used to estimate heritability, define the most heritable phenotype, and explore interactions between genetic and environmental factors. The current diagnostic definitions of major psychiatric disorders are based in part on twin and family data. Epidemiological data are also critical for defining appropriate comparison groups for molecular studies. The data for psychiatric disorders suggest that

TABLE 1. Definition of Terms

Term	Definition
Heritability	Proportion of the variance of a phenotype (disease, trait) that is due to genes, estimated from risks to twins and other relatives.
Mendelian disease	Caused by a (usually rare) change ("mutation") in DNA sequence on one (dominant) or both (recessive) of an individual's pair of chromosomes.
Complex disease	Caused by an interaction of multiple genetic and/or environmental factors.
SNP[a]	Specific position (among 3.2 billion in the genome) where chromosomes carry different nucleic acids. Approximately 11 to 15 million SNPs with frequency ≥1%. Approximately 4 million are catalogued by the HapMap Project.
Common SNPs	≥5% frequency. Approximately 10 million in the genome; approximately 2.8 million on the current HapMap. These SNPs are targeted by GWAS.
Rare variants (rare SNPs)[a]	<1% frequency; many of them very rare. Rarer SNPs in protein-coding regions tend to be more harmful (frequency constrained by selection).
Copy number variant	Chromosomal segment where DNA has been deleted or duplicated. Other structural variants include inversions and translocations.
Common-disease common-variant hypothesis	Some of the genetic risk to common diseases is due to common SNPs.
Multiple rare variant hypothesis	Some of the genetic risk to common disease is due to many different rare SNPs, especially in protein coding or gene regulatory regions.
Linkage disequilibrium between SNPs	Correlation between two SNPs that are close together (an allele of one SNP is usually inherited with a specific allele from the other). Linkage disequilibrium makes GWAS possible: a subset of common SNPs gives information about most of them.
Genomewide association study	A systematic search for common SNPs that influence a disease or trait, using a genomewide SNP array for typing a cohort of individuals. Current arrays also provide information about copy number variants.
Genomewide SNP chip (array)	A system for assaying 300,000 to 1,000,000 SNPs for an individual subject, using an array of bead-based or hybridization assays on a glass slide.

[a] The term "SNP" is sometimes reserved for single-position variants with a frequency ≥1% (i.e., found on at least 1% of chromosomes in a population). For variants with a frequency <1%, the terms "rare variants" and "rare SNPs" are both in use, although "variants" could also refer to other types of sequence changes.

most of the heritable risk is due to interactions of combinations of genetic risk variants, each with a relatively small effect on risk.

Candidate Genes

When the pathophysiology of a disease is known (e.g., an enzyme deficiency), it may be straightforward to define candidate genes and to determine which DNA sequence variants predict who becomes ill. For psychiatric disorders, pathophysiologies are unknown. Most candidate gene hypotheses are based on the effects of psychiatric medications on monoamine neurotransmission, focusing particularly on several functional polymorphisms in dopaminergic or serotonergic pathways (i.e., sequence variants that alter relevant receptor proteins or enzymes) (3, 4). None has been shown to be associated with a psychiatric disorder with a level of significance that would lead to general acceptance of a finding.

Positional Methods

The alternative strategy is to localize disease-related sequence variation based entirely on its location or position in the genome. Before GWAS, available methods included the genomewide linkage study and linkage disequilibrium mapping (of which GWAS is a large-scale example). (See Table 1 for definitions and Table 2 for a timeline of critical developments.)

Genomewide linkage studies became feasible in the 1980s, with genomewide "maps" (7) of hundreds of DNA sequence variations (markers). Linkage analysis (reviewed in [15]), of families with multiple ill members, exploits within-family correlations between illness and the alternative sequences (alleles) of the markers that are closest to the disease-related gene(s). Linkage studies led to the discovery of mutations (mostly rare dominant or recessive) for more than 1,600 diseases (see the Online Mendelian Inheritance in Man [http://www.ncbi.nlm.nih.gov/Omim/mimstats.html]). They have been less successful for complex (multifactorial/multigenic) disorders. In psychiatric linkage studies (catalogued online [https://slep.unc.edu]), small samples of pedigrees were initially examined in the hope of discovering simpler genetic mechanisms that would provide clues to pathophysiology. Then, larger studies (involving hundreds of families) searched for genes with smaller effects. There are diverse opinions regarding the past success and future prospects of these studies. Statistically significant linkages have been reported but have been difficult to replicate, presumably because linkage is much less powerful when risk variants have small effects and there is heterogeneity in the underlying genetic factors in different families. Meta-analyses have supported linkage for some disorders (16–18).

Linkage disequilibrium mapping relies instead on the population-wide correlation between two sequence variants. Most variants are single nucleotide polymorphisms (SNPs) (almost always just two alternative nucleic acids at a genomic position). SNP variants that are reasonably common are mutations that occurred thousands of generations ago and then spread due to chance or natural selection. When a second SNP mutation occurred very close to an earlier one (separated by up to tens of thousands of base pairs), then both variant alleles are almost always transmitted to the same offspring in subsequent generations. Link-

TABLE 2. Timeline of Positional Genetic Methods From Linkage To Genomewide Association Studies

Year	Development	Comment
1980	Proposal to create a genomewide map of DNA markers for human linkage analysis (5).	Following the discovery of restriction fragment length polymorphism markers, it was proposed that once restriction fragment length polymorphisms throughout the genome were available, it would be possible to search any genomic region, or the entire human genome, for evidence of genetic linkage.
1983	Linkage mapping and identification of the Huntington's disease gene (6).	The first of the many Mendelian disorders for which genetic linkage was detected followed by identification of specific disease mutations in the linkage region.
1987	First human linkage map (7).	The first genomewide map of approximately 400 restriction fragment length polymorphisms ushered in the era of genomewide linkage studies. Restriction fragment length polymorphisms were supplanted by short tandem repeat markers and then SNPs.
1993	First genomewide linkage study of a psychiatric disorder (8).	Psychiatric genomewide linkage studies (catalogued online [https://slep.unc.edu]) produced some convergent linkage evidence, but no definitive evidence for susceptibility genes.
1996	Common-disease common-variant hypothesis (9).	The HapMap Project grew out of the need to develop a dense set of genetic markers to test this hypothesis.
2001	Draft of the complete human genome sequence (10).	The genome sequence set the stage for all future progress. It stimulated critical advances in genomic sequencing technology and set a new standard of immediate public release of government-supported genomic research data.
2002–2007	International HapMap Project (11, 12) (www.hapmap.org).	The project discovered and genotyped (in 270 individuals from three populations) 1.3 million SNPs in Phase I plus 2.1 million in Phase II— approximately 25% to 35% of common SNPs in these populations), providing good genomewide coverage. It spurred advances in SNP assays, making genomewide association studies possible. "HapMap III" provided genotypes in an expanded data set for the Illumina 1M and Affymetrix 6.0 (900K) SNP sets.
2002	First published genomewide association study (13).	This study of myocardial infarction used few SNPs (65,761) and cases (94) by current standards.
2005–2007	Availability of high-throughput array-based SNP assays.	Affymetrix and Illumina arrays became available, initially with approximately 100,000 SNPs, and currently with up to approximately 1 million SNPs per array plus additional probes for analysis of copy number. These have made it possible to carry out genomewide association studies for many diseases and samples.
2005	First year with multiple genomewide association study publications.	The first small studies using denser SNP sets produced strong associations for macular degeneration (1) and Crohn's disease (14), demonstrating the feasibility and power of genomewide association studies.
2007	Initiation of the 1,000 Genomes Project (www.1000genomes.org).	This project aims to extend the HapMap to all SNPs with 1% frequency in diverse populations, functional SNPs of lower frequencies, and sequence-level data on structural variants, utilizing multiple high-throughput sequencing technologies.

age disequilibrium is this nonrandom association of two alleles. Approximately 20 years ago, it was proposed that linkage disequilibrium could be exploited to "map" or identify disease genes, such as in linkage candidate regions (or in recently isolated populations in which linkage disequilibrium spans long distances) (19). If one SNP increases the risk of a common disease, then there will be a statistical association in the population between disease and that SNP (direct association) and several nearby SNPs (indirect association due to linkage disequilibrium).

Linkage disequilibrium mapping studies have identified plausible positional candidate genes in regions of linkage or of cytogenetic abnormalities associated with psychiatric disorders, and these genes have suggested new mechanistic hypotheses (20). For example, as of April 2008, there were 1,291 published studies of 690 schizophrenia candidate genes (see http://www.schizophreniaforum.org/res/sczgene/default.asp). A recent meta-analysis of these studies (3) identified four "strong" psychiatric candidate gene associations based on epidemiological criteria for meta-analysis but not at what is currently understood to be a genomewide level of statistical significance.

Common SNPs, HapMap, and GWAS

Risch and Merikangas (21) noted that small genetic effects could be detected with greater power by association analyses and proposed that genomewide linkage disequilibrium mapping (i.e., GWAS) could be applied if technologies were developed to study SNP frequencies in all genes, contrasting ill case subjects versus comparison subjects or case subjects and their parents (associated alleles are transmitted to ill offspring more often than expected by chance). Lander (9) proposed the common-disease common-variant hypothesis. Comparing any two people, most sequence differences are ancient, "common" SNPs (by convention, varying on at least 5% of chromosomes in a population), which Lander argued must confer at least some (not all) of the genetic risk for common diseases. He proposed cataloguing them and studying their association with disease in large samples. SNPs become common because they are neutral or favorable with respect to survival (e.g., evolutionary pressures can rapidly increase frequencies of adaptive SNPs in gene regulating regions). However, some have mildly harmful effects, perhaps depending on environmental conditions (e.g., preserving fat during an ice age but leading to obesity in the fast food era). The common-disease common-variant GWAS strategy assumed that many different common SNPs have small effects on each disease and that some could be found by testing enough SNPs in enough people.

How many SNPs should be tested? Studies of small regions revealed linkage disequilibrium blocks in which

common SNPs are highly correlated (usually <10,000–30,000 base pairs in African populations or 30,000–50,000 base pairs in the newer European and Asian populations) (22). This motivated the HapMap Project (www.hapmap.org [12]), which has validated approximately 4 million SNPs, including 2.8 million of the estimated 10 million common SNPs in major world populations, while creating competition among biotechnology companies to develop high-throughput genotyping technologies. Sequencing and genotyping studies showed that sets of 500,000 (European populations) to 1,000,000 (African populations) SNPs could "tag" (serve as proxies for) approximately 80% of common SNPs (23). Over the last 3 years, the Affymetrix and Illumina companies have developed "chips" (arrays of assays on glass slides) that assay large SNP sets with high accuracy (0%–2% missing data; <0.5% errors), at low cost (approximately $500 [U.S.] per subject; an approximate 2,000-fold reduction in cost per genotype in 10 years), and rapidly (>1,000 DNA specimens per week in some labs). Thus, the GWAS era has arrived.

Rare SNPs

Common SNPs are unlikely to explain all of the genetic risk for common disorders. An evolutionary model of complex diseases (24) predicts roles for common SNPs and for multiple rare variants (such as SNPs) in some genes (multiple rare variant hypothesis). A rare variant is usually defined by a frequency <1%, although many are so rare that they are found in only one individual in a sample (25). Most variants carried by any one person are common, but if one sequences a chromosomal region in many people, one finds more rare SNP sites. The most deleterious variants die out or remain rare due to natural selection (i.e., they reduce survival). They are found in functional regions (i.e., among the SNPs in exons [protein coding regions]) that alter amino acid sequence (nonsynonymous SNPs) or in promoters (sequences that regulate gene expression) (26, 27). However, there are other poorly understood functional regions. Many noncoding regions are highly conserved across species, suggesting that they have a function. Gene expression can be altered by common, synonymous exonic SNPs (no coding change) and by SNPs in introns (noncoding gene segments) (28). Indeed, most genomic DNA is apparently transcribed into RNA and thus could have unknown regulatory functions (29). Most rare SNP associations will be missed by current GWAS methods, but it is expected that the 1,000 Genomes Project (www.1000genomes.org) will discover most SNPs with 1%–5% frequencies, which would permit an extension of systematic GWAS methods to these less common SNPs. Linkage could detect a locus with rare pathogenic variants in many families.

Rare SNP associations are more likely to be detected by resequencing of relevant regions in hundreds or thousands of individuals. (By convention, resequencing, which is sometimes now referred to as "medical sequencing," determines an individual's DNA sequence versus sequencing of an organism's genome.) Botstein and Risch (30) encouraged the systematic study of nonsynonymous SNPs in common diseases. Multiple rare pathogenic variants have been discovered by resequencing genes influencing lipid metabolism (31) and hypertension (32) and also genes for which GWAS have already detected common-SNP associations (33–35). It is anticipated that advances in resequencing technologies will make it feasible to search systematically for rare variant effects in parts of the genome (e.g., linkage regions, all exons, all promoters) and eventually genomewide.

Copy Number Variants

GWAS technologies can also detect more of the copy number variants in the genome than was possible with older cytogenetic methods, by analysis of the relative intensities of the fluorescent labels used in the assays. Copy number variants are deletions and duplications of DNA segments of diverse sizes and population frequencies. For example, large deletions on chromosome 22q11 cause velocardiofacial/DiGeorge syndrome, and 20% of patients with this syndrome also develop schizophrenia (36). Copy number variants tend to arise in regions with repetitive DNA sequences. Some copy number variants are common and transmitted from generation to generation, while others recurrently arise *de novo*. Similar to rare SNPs, rare copy number variants are more likely to be harmful. (Other structural variants, such as inversions and translocations, remain difficult to detect.) Large genomewide copy number variant scans show that copy number variants are more common than previously recognized (37). Structural variation has not been as comprehensively studied as SNPs because copy number variant detection is less accurate, biological confirmation remains costly, and smaller copy number variants (<100,000 base pairs) are less reliably detected. However, technologies are rapidly improving. Significant copy number variant findings are now being reported for psychiatric disorders.

GWAS Study Design

Study design issues are summarized in Table 3. A GWAS sample, selected based on a well-defined heritable phenotype, might include case subjects (ill) and comparison subjects, subjects with a range of values for a continuous phenotypic variable, or probands and both of their parents (trios) or other constellations of relatives. Samples are often limited to a single ancestry (European, Asian, etc.) because some SNPs have markedly different frequencies across populations (and some are not observed in every population) so that some associations can best be detected in homogeneous samples. Each subject is genotyped using a GWAS SNP array. Extensive "quality control" (data cleaning) is required to detect problems that can result in false negative or false positive findings, such as SNPs and DNA specimens that give poor quality results or

TABLE 3. Genomewide Association Study Design Issues and Requirements

Issue	Requirement	Comment
Phenotype	Well-defined, adequately heritable disorder (e.g., schizophrenia) or trait (e.g., high-density cholesterol level or neuroticism score).	Power depends on the frequency and effect size for individual variants, not overall heritability.
Sample type	Ill case subjects and comparison subjects or subjects with a range of trait scores (e.g., highest and lowest) or case subjects and their parents or other relatives.	Case/comparison subjects have more power per subject but are prone to mismatch biases (e.g., ancestry).
Comparison subjects	Match for ancestry, other relevant attributes (e.g., age for an Alzheimer's disease study) or environmental exposures (e.g., "ever smoked" for a study of nicotine dependence).	For more common disorders, comparison subjects with the disorder may be excluded to avoid false negative results (40).
Sample size	Depends on the actual frequency and genetic effect of risk variants in the sample.	Samples up to tens of thousands of subjects have proven useful, but some common risk variants cannot feasibly be detected.
SNPs	300,000 to 1,000,000 common SNPs, depending on ancestry of the sample.	Goal is direct or indirect assay of 80% of HapMap II common SNPs with a correlation (r^2) of ≥0.8.
Multiple testing	P value correction for multiple, partially correlated genotyped SNPs, plus imputed data for all HapMap SNPs to permit cross-study comparison and meta-analysis (40, 41).	Genomewide significance threshold of approximately 5×10^{-8} (42–44).
Population substructure	World populations differ in frequencies of many SNPs. Case subject-comparison subject ancestry differences can create false positive and negative results.	Match case/comparison subjects for ancestry; apply statistical correction for population differences (38).
Data management	Billions of data points to manage.	Requires powerful computers or computer clusters and software (76).
Quality control	Extensive quality control analyses are required to exclude poorly performing SNPs and DNA specimens, identify duplicate or closely-related specimens, and more subtle assay and sample problems.	Without adequate quality control, spurious highly "significant" findings are common.
Detection of copy number variants	Computational methods to detect copy number change from intensities of fluorescent labels in assays; additional nonpolymorphic assays can be added to improve copy number variant detection.	Copy number variant detection is less specific, sensitive, or accurate than SNP genotype detection. Biological confirmation is needed.

unexpected relatedness among subjects. Case-control differences in ancestry ("population substructure") can also confound association test results, but this can be corrected statistically based on correlations among SNP genotypes that reflect ancestry (38). Most studies then test each SNP for association of genotypes to the phenotype and impute the genotypes of other HapMap SNPs based on the correlations among SNPs in HapMap data (39–41).

Selection of comparison groups is critical beyond the problem of ancestral matching. It is ideal to recruit case subjects and comparison subjects systematically from the same population. This is not always feasible for very large samples of a clinically severe disorder, but comparison subjects must be sufficiently comparable with case subjects to avoid systematic biases. Depending on the phenotype, it might be important to match for variables such as age (e.g., for an Alzheimer's disease study) or sex. Information about known gene-environment interactions should be considered (e.g., in studies of substance dependence, comparison subjects are usually selected who have used the substance but did not become dependent). When the phenotype is relatively uncommon (e.g., 5% prevalence), little power is lost by studying comparison subjects without clinical screening, but for more common disorders, power is increased if ill individuals are excluded from the comparison group (40). It is reassuring that in the United Kingdom Wellcome Trust Case Control Consortium GWAS of seven common diseases, robust results were obtained when association was tested using comparison groups recruited from blood donors or from a population-based birth cohort.

Statistical Power of GWAS

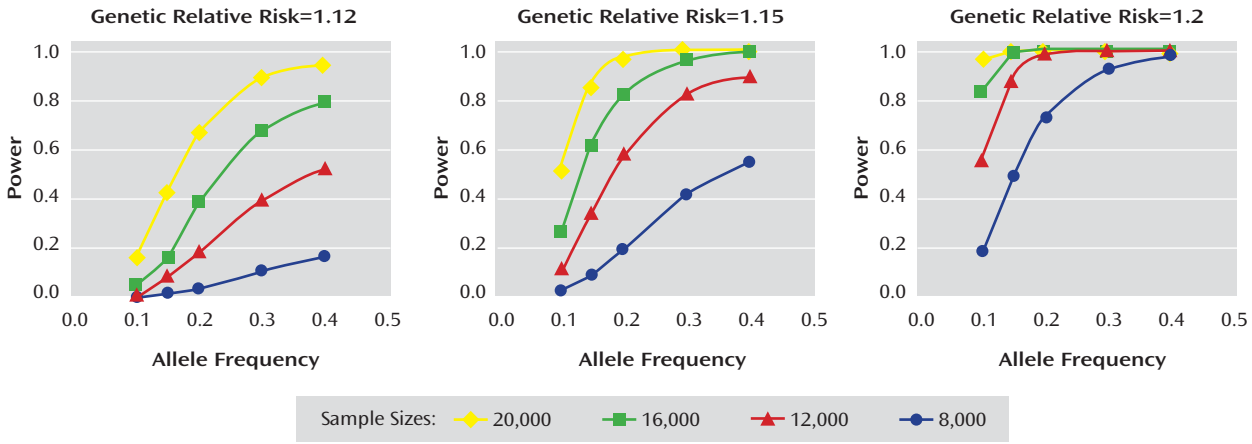
A key factor in the recent success of GWAS has been the assembling of large samples with adequate statistical power to detect small effects of common SNPs on disease risks. Figure 1 illustrates the reasons for this. Large genotypic relative risks (e.g., 5- to 10-fold increase in risk to carriers) would have produced large linkage signals. Early GWAS analyses of several hundred case subjects were powered to search for risk alleles with genotypic relative risks above 2. Only a few such effects were detected (1). The more typical GWAS has included 1,000 to 2,000 case subjects plus a similar number of comparison subjects, with power to detect risk alleles that are reasonably common and have genotypic relative risks of 1.5 to 2. The small number of robust findings suggested the need to detect smaller genotypic relative risks (2).

This led to much larger GWAS analyses in collaborative samples, which have proven to be remarkably successful for many diseases. Most of the new, highly significant findings have been for alleles with genotypic relative risks of 1.1 to 1.4 (mostly between 1.12 and 1.20). In this range (Figure 1), good or excellent power requires samples of 8,000 to 20,000 case subjects (plus comparison subjects), depending on genotypic relative risk and allele frequency (i.e., larger than any sample collected by a single research group to date).

GWAS Findings for Nonpsychiatric Disorders and Lessons for Psychiatry

Over the past 3 years, many highly significant GWAS findings have been reported for nonpsychiatric disorders. Table 4 summarizes a systematic listing of GWAS findings

FIGURE 1. Relationship Among Power, Genotypic Relative Risk (multiplicative inheritance), and Sample Size[a]



[a] The graphs show expected power (91) for a disease with 1% population prevalence ($p=5\times10^{-8}$), depending on minor (less frequent) allele frequency of the tested SNP, sample size (assuming the number of case subjects shown in the graph legend and the same number of comparison subjects, power is similar for the same number of case subject-parent trios), and genotypic relative risk, which is the ratio of the risk of disease to carriers of a particular genotype versus non-carriers (thus, if genotypic relative risk is 1.2, risk is increased by 20%). The calculations assume indirect association between a tested SNP allele and a risk allele at a correlation (r^2) of 0.8, so that the effective sample sizes are approximately 80% of those shown. A sample of 8,000 case subjects and 8,000 comparison subjects will miss most associated alleles that confer much less than a 20% increase in risk (genotypic relative risk <1.2), whereas 20,000/20,000 would detect most associated alleles with genotypic relative risk=1.12 and frequency >15%–20%. Factors that affect power include: 1) **Genotypic relative risk.** Power increases with genotypic relative risk. 2) **Allele frequency and linkage disequilibrium.** Power increases with the minor allele frequency of the associated SNP and with stronger linkage disequilibrium between the SNP and an untested risk allele. 3) **Mode of transmission.** Power is greater for dominant and multiplicative (log additive) genetic effects and less for recessive effects (particularly for rare alleles). 4) Selection of comparison subjects. For diseases with higher prevalence (e.g., >5%), power increases if comparison subjects with the disorder/trait of interest are excluded. (40). 5) **Technical artifacts** of all kinds can reduce power.

provided by the National Institute for Human Genome Research (http://www.genome.gov/GWAstudies/ [accessed Nov. 2008, refs. 42–44]), restricted to findings with p values $<5\times10^{-8}$. There are 200 distinct findings listed for 59 disorders or traits. Some may be false positives due to chance (every p value is an estimate of the probability of a false positive result) or to technical problems such as genotyping or analytical errors. However, many of these findings have already been replicated in independent samples, and most robust p values replicate. These results far exceed all previous robust associations for complex disorders. This confirms that common SNPs explain part of the genetic risk for these disorders, as predicted by the common-disease common-variant hypothesis. There are almost certainly also many common SNPs with smaller effects on risk as well as rare and very rare SNPs and copy number variants with diverse effect sizes.

Sample size. Most initial GWAS samples included 500 to 3,000 case subjects (plus comparison subjects) or as many as 10,657 subjects for a continuous trait. One or more replication samples were usually then studied via collaboration, totaling 2,000 to 8,000 subjects (case and comparison subjects or case subjects and family members). For studies with at least 1,000 case subjects, most findings involved common alleles (20%–80%), with odds ratios (estimates of genotypic relative risk) between 1.1 and 1.4 (i.e., the range within which there was some power).

Findings for type 2 diabetes illustrate the importance of sample size. In late 2007, there were 11 strong candidate genes. Of these, six were discovered by GWAS, four were identified based on mechanistic hypotheses, and one, TCF7L2, was identified by linkage disequilibrium mapping of a linkage region, although TCF7L2 SNPs did not explain the linkage (47). The TCF7L2 gene has an overall odds ratio of 1.37. It was detected by most (not all) studies. Other type 2 diabetes loci have allelic odds ratios between 1.1 and 1.2, requiring from 10,000 to well over 20,000 total subjects for 80% power. Each locus was missed by most single studies. For example, in the Wellcome Trust Case Control Consortium study (2,000 case subjects; 3,000 comparison subjects), these 11 SNPs were ranked from 2 to 26,017 in their strength of association (47). Zeggini et al. (48) combined more than 60,000 subjects to study type 2 diabetes findings that did not quite reach genomewide significance previously. Six SNPs (implicating eight different genes) have since achieved p values $<5\times10^{-8}$, with odds ratios from 1.09 to 1.15.

Novel etiological hypotheses. Most findings have implicated novel genes or regions and suggested new mechanisms. For example, SNPs in FTO ("fat mass and obesity associated" gene) are strongly associated with common obesity (49, 50). This was surprising because FTO knockout mice are not obese. Mechanisms are under study, including a role in adipocyte lipolysis (51). As noted by Todd (52), implicating a gene in disease requires both compel-

TABLE 4. Significant Genomewide Association Study Findings for Nonpsychiatric Disorders[a]

Type of Disease or Trait	Unique Findings With $p \leq 5\times10^{-8}$	Number of Disorders or Traits
Autoimmune	12	3
Bone density	10	1
Cancer	37	8
Cardiovascular	5	4
Diabetes type 1	10	1
Diabetes type 2	10	1
Gastrointestinal	25	5
Lipid levels	13	3
Neurological	9	6
Physical traits	28	7
Plasma values	22	10
Other	19	10
Total	200	59

[a] There is no definitive p value threshold that predicts true positive genomewide association study findings. Interpretation rests on consistency of replication and/or meta-analysis of cumulative data. A p value threshold of 5×10^{-8} has been used throughout this review, based on three estimates that assumed that all common SNPs have been tested (42–44), but other thresholds can be defended. Other approaches include false discovery rate (45) or Bayes Factor (41, 46) criteria. Shown for each category is the number of distinct findings (defined as one or more SNPs in a single chromosomal band for a specific disease or trait) with a p value $\leq5\times10^{-8}$, counting only once those findings reported more than once. Of the 200 findings, 95 had p values $<10^{-12}$, and 58 had p values $<10^{-15}$. Some SNPs or regions have produced findings for different disorders or traits (see article). In many cases, there are additional studies or meta-analyses (not included in this tabulation of genomewide association study reports) that contain additional findings or updated significance levels. "Physical traits" includes nondisease traits such as hair and eye color and height. "Plasma values" includes studies of potentially disease-related values (other than lipids) such as C-reactive protein, glucose, and IgE. "Other" includes studies related to eye, skin, or pulmonary diseases; obesity-related traits; aging; and other traits. Data are summarized online (http://www.genome.gov/GWAstudies/ [accessed November 2008]).

ling statistical evidence for association and substantial additional biological evidence.

Insights into phenotypes. FTO also exemplifies the importance of phenotypic variables. Type 2 diabetes is common in obese individuals. FTO SNPs are associated with type 2 diabetes, but this is due to the association between type 2 diabetes and body mass index (50). The association of FTO with type 2 diabetes is not found when type 2 diabetes case and comparison subjects are matched for body mass index (53). Surprising relationships among phenotypes have also been discovered. For example, SNPs on chromosome 8q24.21 are associated with prostate, breast, and colorectal cancers, which were not previously thought to be genetically related (54). The region contains no known genes, and thus it would have been ignored without a GWAS strategy. It is now being intensively studied.

Thus, GWAS have been remarkably successful for many common diseases. Large multicenter samples have usually been required, and larger samples have detected more associations. Only a small part of the genetic risk for any one disease has been explained, but these discoveries have suggested new disease mechanisms and targets for therapy and prevention, although direct therapeutic applications will require substantial additional effort to characterize the biological mechanisms and develop new treatments. Some of the unexplained variance is likely due to other common SNPs (those that have smaller effects than can be detected with current sample sizes or are not tagged by the arrays or are missed because of technical or sampling problems). The remaining variance may be due to rare SNPs, copy number variants, other unsuspected genomic mechanisms, gene-gene or gene-environment interactions that have not been adequately modeled, and epigenetic effects. The results suggest that the largest possible samples should be studied by GWAS for each of the major psychiatric disorders to test the hypothesis that common SNPs or detectable copy number variants are involved in etiology. Positive findings could lead to important etiological discoveries.

GWAS of Psychiatric Disorders

GWAS findings are now emerging for psychiatric disorders (Table 5). The early findings include replicated copy number variant associations for schizophrenia and autism, a genomewide significant association for bipolar disorder that emerged when several data sets were combined, and a significant association in a combined schizophrenia-bipolar data set.

For schizophrenia, four genomewide studies of copy number variants (55–58) have produced two types of replicated findings. First, two large studies (55, 56) found two rare deletions that are significantly associated with schizophrenia on chromosomes 1q21.1 (0.2% of case subjects) and 15q13.3 (0.3% of case subjects). The case: comparison subjects ratio (approximately 10) suggests major effects on risk, but it is unknown which deleted genes or sequences are responsible or whether they account for all of the subjects' genetic risks. These deletions are also found (but probably less frequently) in individuals with mental retardation and/or autism and are typically *de novo* (not inherited from parents) (55). The well-known chromosome 22q11 deletions were also significantly associated with schizophrenia (0.2%–0.4% of case subjects across studies versus 0% of comparison subjects). Second, the three studies (56–58) that tested such a hypothesis showed that schizophrenia case subjects have a small but significant increase in their total genomewide count of rare long copy number variants, suggesting that there are other patho-

TABLE 5. Published Genomewide Association Studies of Psychiatric Disorders[a]

Item	Disorder	Initial Sample (case subjects/ comparison subjects)	Other Data	Genomewide Significant Findings
Studies of association to SNP genotypes (individual genotyping)				
Wellcome Trust Case Control Consortium (41)	Bipolar disorder	1,868/2,938 (United Kingdom)		
Sklar et al. (59)	Bipolar disorder	1,461/2,008 (United States, United Kingdom, Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)-University College of London)	Replication sample (for best results): 409 U.S. trios; 365 case subjects/351 comparison subjects (Scottish)	
Ferreira et al. (60)	Bipolar disorder	4,387/6,209	Wellcome Trust Case Control Consortium (41) plus Sklar (59) (see numbers above) plus ED-DUB-STEP2 (1,098 case subjects/ 1,267 comparison subjects)	$p=9.1\times10^{-9}$ ANK3 gene (odds ratio=1.45; frequency of the associated allele in comparison subjects=0.053)
Lencz et al. (61)	Schizophrenia	178/144 (United States)		
Sullivan et al. (62)	Schizophrenia	738/733 (United States)	Multiple ancestries	
O'Donovan et al. (63)	Schizophrenia	479/2,937 (United Kingdom) plus 1,865 Wellcome Trust Case Control Consortium bipolar disorder case subjects	Replication sample (for best results): 6,829 case subjects/9,897 comparison subjects (United Kingdom, Europe, United States, Australia, Japan, Israel	With bipolar disorder included: $p=9.96\times10^{-9}$; ZNF804A gene (odds ratio=1.12; frequency of the associated allele in comparison subjects=0.59)
Studies of association to copy number variants				
Walsh et al. (57)	Schizophrenia	150/268	Replication sample (for best results): 83 childhood-onset schizophrenia patients plus parents	p=0.0008; increased novel copy number variants in case subjects (15%) versus comparison subjects (5%) (p=0.03 in child-onset schizophrenia)
Xu et al. (58)	Schizophrenia	152/159	Sporadic case subjects	p=0.0008; increased noninherited copy number variants in sporadic case subjects (9.9%) versus comparison subjects (1.26%)
Stone et al. (56)	Schizophrenia	3,381/3,191		$p=3\times10^{-5}$; increased copy number variants (<1% frequency; >100Kb) in case subjects (1.14% per subject) versus comparison subjects (0.99). Genomewide significant evidence for association of copy number variants on chromosome 1q21.1, 22q11.2, 15q13.3.
Stefansson et al. (55)	Schizophrenia	1,433/33,350	Replication sample (for best results): 3,285 case subjects/ 7,951 comparison subjects	Genomewide significant evidence for association of copy number variants on 1q21.1, 22q11.2, 15q11.2, 15q13.3
Sebat et al. (64)	Autism	118 (sporadic)/196	Some comparison subjects from autism families; some Autism Genetics Resource Exchange families	Increased *de novo* copy number variants in case subjects (10%) versus comparison subjects (1%); (note: >1 comparison subject per family)
Kumar 2008 (65)	Autism	180/372	Replication sample (for best results): 532 case subjects/465 comparison subjects	p=0.044 (uncorrected); increased 16p.11.2 deletions in case subjects (0.6%) versus comparison subjects (0%)
Marshall et al. (66)	Autism	427/500	Replication sample (for best results): 1,152 additional comparison subjects	Genomewide significant evidence for association of increased *de novo* copy number variants in case subjects (7%) versus comparison subjects (1%). Increased 16p.11.2 deletions in case subjects (approximately 1%) versus comparison subjects (0%); (p=0.002)
Weiss et al. (67)	Autism	751 multiplex Autism Genetics Resource Exchange families (1,441 case subjects) plus 2,814 comparison subjects	Replication samples (for best results): 512 case subjects/434 comparison subjects; 299 case subjects/18,834 comparison subjects	Increased 16p.11.2 copy number variants in case subjects (1.1%) versus comparison subjects (0.05%); significant in all three samples
Christian et al. (68)	Autism	397/372	Case subjects from Autism Genetics Resource Exchange families	11.6% of case subjects had a copy number variant unique to case subjects

[a] The copy number variant studies of Sebat et al., Weiss et al., and Christian et al. all used some families from the Autism Genetics Resource Exchange repository and thus are not entirely independent. SNP studies all used SNP arrays with 500,000 SNPs (Affymetrix 500K or 5.0) or 900,000 SNPs (Affymetrix 6.0). Some studies used more than one type of array. Copy number variant studies used array-based comparative genomic hybridization and/or genomewide association study SNP arrays (Affymetrix or Illumina), with additional confirmation of some or all results using additional methods (e.g., quantitative polymerase chain reaction, karyotyping, and other methods). Studies using pooled genotyping are not included but are cited in the article text.

genic copy number variants that are too rare to detect singly. Other GWAS studies of psychiatric disorders are summarized in Table 5.

At the time of this writing, there have been three published small schizophrenia GWAS analyses using individual genotyping (61–63), with samples that included 178 to 738 case subjects. Two additional studies (69, 70) used pooled genotyping. To date, no genomewide significant finding has emerged for schizophrenia alone. However, when the 12 "best" SNPs from a GWAS of 479 case subjects and 2,937 Wellcome Trust Case Control Consortium comparison subjects were genotyped in an additional 7,308 schizophrenia case subjects and 12,834 comparison subjects, and the 1,868 Wellcome Trust Case Control Consortium bipolar disorder case subjects were added to the analysis, a genomewide significant p value was seen for a SNP in a gene of unknown function (zinc finger protein 804A [ZNF804A]) (63). This will require replication for these disorders, both separately and combined. It illustrates the potential importance of cross-diagnosis analyses, although such analyses will increase the problem of multiple testing and thus require very large samples for confirmation.

For autism, three studies (65–67) have reported association with a rare (1% of case subjects), large, high-penetrance deletion on chromosome 16p11.2. There is also support for the hypothesis that there is an excess of rare, mostly *de novo*, copy number variants in approximately 10% of case subjects, although the role of these rare copy number variants in autism remains to be proven (64, 65, 68). Autism GWAS analyses of common SNPs have yet to be reported.

For bipolar disorder, three individual studies (41, 59, 60), of 1,000 to 2,000 case subjects each, failed to detect significant association, but the three data sets combined produced a p value of 9.1×10^{-9} for ankyrin-G (ANK3), the product of which links membrane proteins, such as voltage-dependent sodium channels, to the axonal cytoskeleton. The larger analysis did not replicate a significant association (for DGKH [diacylglycerol kinase, eta]) that had been reported in a smaller study using pooled genotyping (71).

Among the reports that will be available in the near future are the four psychiatric GWAS reports, supported by the Genetic Association Information Network (fnih.org), on schizophrenia, bipolar disorder, major depression, and ADHD. Details and preliminary results are available online (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap). We are not permitted to summarize these analyses pending the initial publications by the investigators. The Genetic Association Information Network is an example of a new emphasis on rapid public sharing of genetic data to accelerate the process of discovery.

The Psychiatric GWAS Consortium

The first set of psychiatric GWAS analyses have demonstrated that this methodology can work for psychiatric disorders. The pattern observed in the bipolar disorder studies is particularly encouraging because it is consistent with what has happened for nonpsychiatric diseases: combining several smaller samples produced a significant result as well as several other findings with modestly significant p values in each individual study, which could prove to be significant as more data become available (60).

These results support our expectation that multiple definitive association findings will be detected for many psychiatric disorders, often requiring large samples. We therefore organized the Psychiatric GWAS Consortium, which includes almost all known GWAS studies to date for schizophrenia, bipolar disorder, major depressive disorder, ADHD, and autism (Table 6), contributed by 121 investigators from 61 institutions at the time of this writing (as listed in the Acknowledgments). The Psychiatric GWAS Consortium has the following three specific objectives:

1. *Within-disorder meta-analyses of all available GWAS data.* These diagnoses are based on definitions that produced maximum heritability estimates in genetic epidemiological studies. Thus, disorder-specific analyses represent our strongest hypotheses.
2. *Cross-disorder analyses*, including analyses of combinations of disorders and of phenotypes observed in two or more disorders (such as depression or psychosis) based on the recommendations of an expert committee. Because data are insufficient to determine what common cross-disorder etiological factors might exist, alternative phenotypes should be explored. GWAS analyses have produced surprising cross-disorder associations, such as those found for cancers (54) and inflammatory bowel diseases (86), which could also exist for psychiatric disorders given the many common symptoms.
3. *Analyses of comorbidities* such as alcohol, nicotine, and illicit drug use disorders, which disorders can be studied across multiple case groups.

Additional exploratory analyses will be carried out by analysts from participating research groups, generating new hypotheses that can be tested as more samples become available. All GWAS data used by the Psychiatric GWAS Consortium will become available to the scientific community through data repositories (unless prohibited by the original consents or institutional review board decisions).

A central analytical team in consultation with participating analysts will conduct uniform quality control analyses and imputation of nongenotyped HapMap SNPs (to permit combining of data). The disorder-specific workgroups will design their own primary meta-analyses, with additional workgroups to define other phenotypic and cross-disorder analyses. Analyses will account for ethnic substructure within samples and appropriate pairing of case and comparison groups.

Depending on the genetic architecture of each disorder, one or more primary analyses could have sufficient power to detect genomewide significant evidence for associa-

TABLE 6. Summary of Psychiatric GWAS Consortium Genomewide Association Study Samples and Characteristics of Studied Disorders[a]

Disorder	Samples	Case Subjects	Comparison Subjects	Trios or Families	Prevalence Rate (%)	Heritability (%)
ADHD	6	1,418	0	2,443	4–12	70–80
Autism	6	652	6,000	4,661	0.3–0.6	90–100
Bipolar disorder	10	7,075	10,559	0	0.3–1.5	73–93
Major depressive disorder	9	12,926	9,618	0	5–18	31–42[b]
Schizophrenia	11	9,588	13,500	650	0.2–1.1	73–90
Total	42	31,659	26,945	7,772		

[a] Data shown are expected combined sample sizes for meta-analysis of genomewide association study data by the Psychiatric GWAS Consortium by the end of 2009. Data are reported for subjects of European-ancestry only; a small number of African American samples are also available for schizophrenia and bipolar disorder. The case subjects are all independent (although independence is tested using genotypes). For each disorder, comparison subjects used in more than one study are counted once; also, comparison subjects used for more than one disorder are counted once in the Total, which is therefore less than the sum of the rows for the disorders. The column for "Trios or Families" includes a sample of multiply-affected families for schizophrenia and trio or sib-pair families (with parents) for ADHD and autism. References for prevalence and heritability are as follows: ADHD (72–74), autism (75, 76), bipolar disorder (77, 78), major depressive disorder (79–82), and schizophrenia (83, 84).

[b] For major depression, higher estimates have been obtained in clinical samples (85) or using repeated interviews (81).

tion. For example, the largest analyses, with approximately 10,000 case subjects and 10,000 comparison subjects, would have 80% power to detect a SNP with a genotypic relative risk of 1.152 and p value $<5\times10^{-8}$—assuming direct association with an allele with a frequency of 0.25 and log-additive inheritance—or 57% power for indirect association with an r^2 value of 0.8. Power would be reduced for smaller samples, less common alleles, and recessive effects. If there are many risk alleles in the genome with a sufficient effect size, there would be substantial power to detect at least one. Meta-analyses will be completed in 2009. Updated results will be made available on the Psychiatric GWAS Consortium website (http://pgc.unc.edu).

Discussion

There is a compelling rationale for applying GWAS methods to very large samples for major psychiatric disorders. Given that the pathophysiologies of these disorders are unknown, genomewide studies provide an unbiased way to search the genome for causative factors. Many successful GWAS analyses have combined data from diverse clinical samples and SNP arrays to obtain replicable findings that point to new hypotheses about disease mechanisms and treatment targets. The first significant psychiatric GWAS findings have been reported (Table 5), using large collaborative samples. It is hoped that meta-analyses can produce multiple robust findings for psychiatric disorders.

GWAS SNP arrays "cover" ≥80% of common HapMap SNPs, and regional resequencing data suggest that most unknown common SNPs are also being tested indirectly. Within these limitations, GWAS methods test the common-disease common-variant hypothesis. Copy number variants are also detected but less systematically or accurately. The Psychiatric GWAS Consortium meta-analyses will have reasonable power to detect common SNP associations for each disorder within the limitations illustrated in Figure 1. However, it is possible that very few significant associations might be detected for some disorders or none. How far should we go with GWAS?

Past experience suggests that for some disorders as many as 20,000 to 30,000 case subjects and a similar number of comparison subjects (or case subjects plus their parents) could be required to obtain highly robust findings. More data sets will be genotyped in the near future, and the National Institute of Mental Health plans to collect additional large schizophrenia and bipolar disorder samples (http://grants.nih.gov/grants/guide/rfa-files/RFA-MH-08-131.html). This raises important questions regarding resource allocation. For example, the next phase of genetic studies will involve a combination of increasingly large GWAS analyses (for common SNP and copy number variant associations) and resequencing studies (for rare variants). It is not known how these and other research investments should be optimally balanced.

To the extent that resources are available, we encourage a long-term view, avoiding the well-known pattern of initial exuberance followed by disillusionment. The logic of GWAS has been clear for more than 10 years (23). Results have been remarkably consistent with expectations in the sense that common SNP associations have been discovered for many common disorders, particularly those that have been studied with larger sample sizes. It is true that initial GWAS results have explained only a small part of the etiological variance for each disease, and it seems certain that studies of copy number variants and rare SNPs will also be critical in elucidating disease mechanisms. However, it is likely that common SNPs explain a larger portion of the variance than that which can be determined with existing sample sizes, with many common SNPs, each with small effects, contributing collectively to a major portion of genetic risk (24). As the number of associations increases, the biological pathways underlying risk for each disease become clearer. GWAS methods should be applied systematically to major psychiatric disorders in large samples. The following caveats are important to consider:

1. Some disorders might not be amenable to GWAS (e.g., if all risk alleles have very low genotypic relative risks or if genetic risks are conferred by multiple rare SNPs or by copy number variants too small to be detected reliably). Discoveries for these disorders might only be possible with larger-scale resequencing studies.
2. Current diagnostic categories might be inadequate. Endophenotypic variables (neuroimaging, electrophysiological, neuropsychological, biochemical, and other markers) might better index the underlying gene effects (87), although none has yet proven more heritable than diagnostic categories. These measures are not usually available in large data sets.
3. Genetic heterogeneity reduces power. A low frequency risk allele is an example of heterogeneity (i.e., most case subjects do not share that risk factor). Power (Figure 1) is best for frequencies above approximately 20% and poor for frequencies greatly below 10% unless genotypic relative risk is high. Heterogeneity might be increased in large multicenter samples. For example, despite the generally high interrater reliability for these disorders, research groups can have diagnostic "biases," some of which could correlate with specific risk alleles. However, power increases with sample size despite some degree of misclassification, and classification is imperfect for many medical disorders for which there are GWAS findings.
4. More needs to be learned about the selection of comparison subjects for psychiatric GWAS studies. It remains possible that some findings will be confounded by systematic biases in comparison groups, such as under-representation of developmental disabilities. The field will need much larger comparison groups ascertained by diverse methods and from multiple ethnic populations.
5. For some disorders there might be no detectable main effects of SNPs, only higher order gene-gene or gene-environment interactions. However, main effects are often detectable even if interactions are erroneously excluded. Explicit tests of interactions (88) or data mining might prove informative.
6. GWAS assays do not interrogate all common variants. For each array type, some assays perform poorly and some common SNPs are not or cannot be tagged.
7. Improved methods will be needed to provide more systematic information about copy number variants and their relationship to disease. Associated copy number variant regions will require resequencing studies of large numbers of subjects without copy number variants to determine whether these regions also contain rare, highly penetrant associated variants.
8. There are probably unknown genetic mechanisms. We have only recently recognized the importance of copy number variants, micro RNAs, long-range promoters, and epigenetic factors (genomic effects other than sequence changes such as DNA methylation patterns) (89). The discovery that most of the genome is transcribed suggests that many types of functional sequence are undiscovered (12).

Bearing these risks and caveats in mind, we conclude that GWAS methods have discovered a remarkable set of robust common SNP association findings for a broad range of diseases (90), now including an initial set of SNP and copy number variant associations for psychiatric disorders. It is reasonable to predict that studies of sufficiently large samples can produce definitive discoveries of genetic risk factors for psychiatric disorders and that these discoveries will contribute to the definitive identification of pathophysiological mechanisms for the first time.

This article was written by the Psychiatric GWAS Consortium Coordinating Committee, whose members (presented in alphabetical order) take responsibility for its content: Sven Cichon, Ph.D. (University of Bonn, Germany); Nick Craddock, M.D., Ph.D. (Cardiff University); Mark Daly, Ph.D. (Harvard Medical School, Broad Institute); Stephen V. Faraone, Ph.D. (State University of New York Upstate Medical University); Pablo V. Gejman, M.D. (Northshore University HealthSystem and Feinberg School of Medicine of Northwestern University); John Kelsoe, M.D. (University of California, San Diego); Thomas Lehner, Ph.D., M.P.H. (NIMH); Douglas F. Levinson, M.D. (Stanford University); Audra Moran, M.A. (NARSAD, Ex Officio); Pamela Sklar, M.D., Ph.D. (Massachusetts General Hospital, Broad Institute); and Patrick F. Sullivan, M.D. (University of North Carolina at Chapel Hill). Received Sept. 11, 2008; revision received Nov. 17, 2008; accepted Nov. 20, 2008 (doi: 10.1176/appi.ajp.2008.08091354). From the Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, Calif. Address correspondence and reprint requests to Dr. Levinson, Department of Psychiatry and Behavioral Sciences, Stanford University, 710 Welch Rd., Suite A-3325, Palo Alto, CA 94304; dflev@stanford.edu (e-mail).

Dr. Faraone receives research support from or has served on the advisory boards of Shire, Eli Lilly, Pfizer, McNeil, and NIH. Dr. Kelsoe is a founder of and holds equity in Psynomics, Inc. Dr. Sullivan has received unrestricted research support from Eli Lilly for genetic research in schizophrenia. Drs. Cichon, Craddock, Daly, Gejman, Lehner, Levinson, Sklar, and Sullivan and Ms. Moran report no competing interests.

Supported by NIMH grant MH-085520. Statistical analyses were conducted using the Genetic Cluster Computer, which is supported by the Netherlands Scientific Organization (NWO 480–05-003, PI Danielle Posthuma), along with a supplement from the Dutch Brain Foundation.

ADHD Working Group: Stephen Faraone, Chair (SUNY-UMU); Richard Anney (Trinity College Dublin); Jan Buitelaar (Radboud University); Josephine Elia (Children's Hospital of Philadelphia); Barbara Franke (Radboud University); Michael Gill (Trinity College Dublin); Hakon Hakonarson (CHOP); Lindsey Kent (St. Andrews University); James McGough (UCLA); Eric Mick (Massachusetts General Hospital/Harvard University); Laura Nisenbaum (Eli Lilly); Susan Smalley (UCLA); Anita Thapar (Cardiff University); Richard Todd, deceased (Washington University/St. Louis, MO); and Alexandre Todorov (Washington University/St. Louis, MO).

Autism Working Group: Bernie Devlin, Chair (University of Pittsburgh); Mark Daly, Co-Chair (Massachusetts General Hospital/Harvard University); Richard Anney (Trinity College Dublin); Dan Arking (Johns Hopkins University); Joseph D. Buxbaum (Mt. Sinai School of Medicine, New York); Aravinda Chakravarti (Johns Hopkins University); Edwin Cook (University of Illinois); Michael Gill (Trinity College Dublin); Leena Peltonen (University of Helsinki); Joseph Piven (University of North Carolina-Chapel Hill); Guy Rouleau (University of Montreal); Susan Santangelo (Massachusetts General Hospital/Harvard University); Gerard Schellenberg (University of Washington); Steve Scherer (University of Toronto); James Sutcliffe (Vanderbilt University); Peter Szat-

mari (McMaster University); and Veronica Vieland (Columbus Children's Research Institute).

Bipolar Disorder Working Group: John Kelsoe, Co-Chair (UCSD); Pamela Sklar, Co-Chair, (Massachusetts General Hospital/Harvard University); Ole A. Andreassen (University of Oslo, Norway); Douglas Blackwood (University of Edinburgh, Scotland); Michael Boehnke (University of Michigan); Rene Breuer (CIMH, Mannheim, Germany); Margit Burmeister (University of Michigan); Sven Cichon (University of Bonn, Germany); Aiden Corvin (Trinity College Dublin); Nicholas Craddock (Cardiff University); Manuel Ferreira (Massachusetts General Hospital/Harvard University); Matthew Flickinger (University of Michigan); Tiffany Greenwood (UCSD); Weihua Guan (University of Michigan); Hugh Gurling (University College London); Jun Li (University of Michigan); Eric Mick (Massachusetts General Hospital/Harvard University); Valentina Moskvina (Cardiff University); Pierandrea Muglia (GlaxoSmithKline); Walter Muir (University of Edinburgh, Scotland); Markus Noethen (University of Bonn, Germany); John Nurnberger (Indiana University); Shaun Purcell (Massachusetts General Hospital/Harvard University); Marcella Rietschel (CIMH, Mannheim); Douglas Ruderfer (Massachusetts General Hospital/Harvard University); Nicholas Schork (UCSD); Thomas Schulze (CIMH, Mannheim); Laura Scott (University of Michigan); Michael Steffens (University of Bonn, Germany); Ruchi Upmanyu (GlaxoSmithKline); and Thomas Wienker (University of Bonn, Germany).

Cross-Disorder Working Group: Jordan Smoller, Co-Chair (Massachusetts General Hospital/Harvard University); Nicholas Craddock, Co-Chair (Cardiff University); Kenneth Kendler, Co-Chair (Virginia Commonwealth University); John Nurnberger (Indiana University); Roy Perlis (Massachusetts General Hospital/Harvard University); Shaun Purcell (Massachusetts General Hospital/Harvard University); Marcella Rietschel (CIMH, Mannheim); Susan Santangelo (Massachusetts General Hospital/Harvard University); and Anita Thapar (Cardiff University).

Major Depressive Disorder Working Group: Patrick Sullivan, Chair (University of North Carolina-Chapel Hill); Douglas Blackwood (University of Edinburgh, Scotland); Dorret Boomsma (Vrije University, Amsterdam); Rene Breuer (CIMH, Mannheim, Germany); Sven Cichon (University of Bonn, Germany); William Coryell (University of Iowa); Eco de Geus (Vrije University, Amsterdam); Steve Hamilton (UCSF); Witte Hoogendijk (Vrije University, Amsterdam); Stafam Kloiber (MPI-P Munich); William B. Lawson (Howard University); Douglas Levinson (Stanford University); Cathryn Lewis (IOP, London); Susanne Lucae (MPI-P Munich); Nick Martin (QIMR); Patrick McGrath (Columbia University); Peter McGuffin (IOP, London); Pierandrea Muglia (GlaxoSmithKline); Walter Muir (University of Edinburgh, Scotland); Markus Noethen (University of Bonn, Germany); James Offord (Pfizer); Brenda Penninx (Vrije University, Amsterdam); James B. Potash (Johns Hopkins University); Marcella Rietschel (CIMH, Mannheim, Germany); William A. Scheftner (Rush University); Thomas Schulze (CIMH, Mannheim); Susan Slager (Mayo Clinic); Federica Tozzi (GlaxoSmithKline); Myrna M. Weissman (Columbia University); AHM Willemsen (Vrije University, Amsterdam); and Naomi Wray (QIMR).

Schizophrenia Working Group: Pablo Gejman, Chair (Northshore University HealthSystem and Feinberg School of Medicine of Northwestern University); Ole A. Andreassen (University of Oslo, Norway); Douglas Blackwood (University of Edinburgh, Scotland); Sven Cichon (University of Bonn, Germany); Aiden Corvin (Trinity College Dublin); Mark Daly (Massachusetts General Hospital/Harvard University); Ayman Fanous (Washington Veterans Administration Medical Center, Georgetown University, Virginia Commonwealth University); Michael Gill (Trinity College Dublin); Hugh Gurling (UCL); Peter Holmans (Cardiff University); Christina Hultman (Karolinska Institutet); Kenneth Kendler (Virginia Commonwealth University); Sari Kivikko (National Public Health Institute); Claudine Laurent (Pierre and Marie Curie Faculty of Medicine, Paris); Todd Lencz (LIJ); Douglas Levinson (Stanford University); Anil Malhotra (LIJ); Bryan Mowry (Queensland Center for Mental Health Research, University of Queensland); Markus Noethen (University of Bonn, Germany); Mike O'Donovan (Cardiff University); Roel Ophoff (UCLA); Michael Owen (Cardiff University); Leena Peltonen (University of Helsinki); Ann Pulver (Johns Hopkins University); Marcella Rietschel (CIMH, Mannheim); Brien Riley (Virginia Commonwealth University); Alan Sanders (Northshore University HealthSystem and Feinberg School of Medicine of Northwestern University); Thomas Schulze (CIMH, Mannheim); Sibylle Schwab (University of Western Australia); Pamela Sklar (Massachusetts General Hospital/Harvard University); David St. Clair (University of Aberdeen); Patrick Sullivan (University of North Carolina-Chapel Hill); Jaana Suvisaari (University of Helsinki); Edwin van den Oord (Virginia Commonwealth University); Naomi Wray (QiMR); and Dieter Wildenaver (Univerisity of Western Australia).

Statistical Analysis and Computational Working Group: Mark Daly, Chair (Massachusetts General Hospital/Harvard University); Phillip Awadalla (University of Montreal); Bernie Devlin (University of Pittsburgh); Frank Dudbridge (MRC-BSU); Arnoldo Frigessi (University of Oslo, Norway); Elizabeth Holliday (QCMHR/University of Queensland); Peter Holmans (Cardiff University); Todd Lencz (LIJ), Douglas Levinson (Stanford University); Cathryn Lewis (IOP, London); Danyu Lin (University of North Carolina-Cahpel Hill); Valentina Moskvina (Cardiff University); Bryan Mowry (QCMHR/University of Queensland); Ben Neale (Massachusetts General Hospital/Harvard University), Eve Pickering (Pfizer Pharmaceuticals Group); Danielle Posthuma (Vrije University Amsterdam); Shaun Purcell (Massachusetts General Hospital/Harvard University); John Rice (Washington University/St. Louis, MO); Stephan Ripke (MPI-P Munich); Nicholas Schork (UCSD); Jonathan Sebat (CSHL); Michael Steffens (University of Bonn, Germany); Jennifer Stone (Massachusetts General Hospital/Harvard University); Jung-Ying Tzeng (NCSU); Edwin van den Oord (Virginia Commonwealth University); and Veronica Vieland (Columbus Children's Research Institute).

The authors thank the Psychiatric GWAS Consortium colleagues for their contributions. The authors also thank NARSAD for infrastructure support.

References

1. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J: Complement factor H polymorphism in age-related macular degeneration. Science 2005; 308:385–389
2. Manolio TA, Brooks LD, Collins FS: A HapMap harvest of insights into the genetics of common disease. J Clin Invest 2008; 118: 1590–1605
3. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ, Tanzi RE, Bertram L: Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the Szgene database. Nat Genet 2008; 40:827–834
4. Levinson DF: Meta-analysis in psychiatric genetics. Curr Psychiatry Rep 2005; 7:143–151
5. Botstein D, White RL, Skolnick M, Davis RW: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 1980; 32:314–331
6. Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi AY: A polymorphic DNA marker genetically linked to Huntington's disease. Nature 1983; 306:234–238
7. Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, Bowden DW, Smith DR, Lander ES, Botstein D, Akots G, Rediker KS, Gravius T, Brown VA, Rising MB, Parker C, Powers JA, Watt DE, Kauffman ER, Bricker A, Phipps P, Muller-Kahle H, Fulton TR, Ng S, Schumm JW, Braman JC, Knowlton RG, Barker DF, Crooks SM, Lincoln SE, Daly MJ, Abrahamson J: A genetic linkage map of the human genome. Cell 1987;51:319–337
8. Coon H, Jensen S, Hoff M, Holik J, Plaetke R, Reimherr F, Wender P, Leppert M, Byerley W: A genome-wide search for genes predisposing to manic-depression, assuming autosomal dominant inheritance. Am J Hum Genet 1993; 52:1234–1249
9. Lander ES: The new genomics: global views of biology. Science 1996; 274:536–539
10. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Mor-

ris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C: Initial sequencing and analysis of the human genome. Nature 2001; 409:860–921

11. Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu FL, Yang HM, Ch'ang LY, Huang W, Liu B, Shen Y, Tam PKH, Tsui LC, Waye MMY, Wong JTF, Zeng CQ, Zhang QR, Chee MS, Galver LM, Kruglyak S, Murray SS, Oliphant AR, Montpetit A, Hudson TJ, Chagnon F, Ferretti V, Leboeuf M, Phillips MS, Verner A, Kwok PY, Duan SH, Lind DL, Miller RD, Rice JP, Saccone NL, Taillon-Miller P, Xiao M, Nakamura Y, Sekine A, Sorimachi K, Tanaka T, Tanaka Y, Tsunoda T, Yoshino E, Bentley DR, Deloukas P, Hunt S, Powell D, Altshuler D, Gabriel SB, Qiu RZ, Ken A, Dunston GM, Kato K, Niikawa N, Knoppers BM, Foster MW, Clayton EW, Wang VO, Watkin J, Gibbs RA, Belmont JW, Sodergren E, Weinstock GM, Wilson RK, Fulton LL, Rogers J, Birren BW, Han H, Wang HG, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Todani K, Fujita T, Tanaka S, Holden AL, Lai EH, Collins FS, Brooks LD, McEwen JE, Guyer MS, Jordan E, Peterson JL, Spiegel J, Sung LM, Zacharia LF, Kennedy K, Dunn MG, Seabrook R, Shillito M, Skene B, Stewart JG, Valle DL, Clayton EW, Jorde LB, Belmont JW, Chakravarti A, Cho MK, Duster T, Foster MW, Jasperse M, Knoppers BM, Kwok PY, Licinio J, Long JC, Marshall PA, Ossorio PN, Wang VO, Rotimi CN, Royal CDM, Spallone P, Terry SF, Lander ES, Lai EH, Nickerson DA, Abecasis GR, Altshuler D, Bentley DR, Boehnke M, Cardon LR, Daly MJ, Deloukas P, Douglas JA, Gabriel SB, Hudson RR, Hudson TJ, Kruglyak L, Kwok PY, Nakamura Y, Nussbaum RL, Royal CDM, Schaffner SF, Sherry ST, Stein LD, Tanaka T: The International HapMap Project. Nature 2003; 426:789–796

12. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J: A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449: 851–861

13. Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Sato H, Hori M, Nakamura Y, Tanaka T: Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet 2002; 32:650–654

14. Yamazaki K, McGovern D, Ragoussis J, Paolucci M, Butler H, Jewell D, Cardon L, Takazoe M, Tanaka T, Ichimori T, Saito S, Sekine A, Iida A, Takahashi A, Tsunoda T, Lathrop M, Nakamura Y: Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum Mol Genet 2005; 14:3499–3506

15. Teare MD, Barrett JH: Genetic linkage studies. Lancet 2005; 366:1036–1044

16. Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, Williams NM, Schwab SG, Pulver AE, Faraone SV, Brzustowicz LM, Kaufmann CA, Garver DL, Gurling HM, Lindholm E, Coon H, Moises HW, Byerley W, Shaw SH, Mesen A, Sherrington R, O'Neill FA, Walsh D, Kendler KS, Ekelund J, Paunio T, Lonnqvist J, Peltonen L, O'Donovan MC, Owen MJ, Wildenauer DB, Maier W, Nestadt G, Blouin JL, Antonarakis SE, Mowry BJ, Silverman JM, Crowe RR, Cloninger CR, Tsuang MT, Malaspina D, Harkavy-Friedman JM, Svrakic DM, Bassett AS, Holcomb J, Kalsi G, McQuillin A, Brynjolfson J, Sigmundsson T, Petursson H, Jazin E, Zoega T, Helgason T: Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: schizophrenia. Am J Hum Genet 2003; 73:34–48

17. McQueen MB, Devlin B, Faraone SV, Nimgaonkar VL, Sklar P, Smoller JW, Abou Jamra R, Albus M, Bacanu SA, Baron M, Barrett TB, Berrettini W, Blacker D, Byerley W, Cichon S, Coryell W, Craddock N, Daly MJ, Depaulo JR, Edenberg HJ, Foroud T, Gill M, Gilliam TC, Hamshere M, Jones I, Jones L, Juo SH, Kelsoe JR,

Lambert D, Lange C, Lerer B, Liu J, Maier W, Mackinnon JD, McInnis MG, McMahon FJ, Murphy DL, Nothen MM, Nurnberger JI, Pato CN, Pato MT, Potash JB, Propping P, Pulver AE, Rice JP, Rietschel M, Scheftner W, Schumacher J, Segurado R, Van Steen K, Xie W, Zandi PP, Laird NM: Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q. Am J Hum Genet 2005; 77:582–595

18. Trikalinos TA, Karvouni A, Zintzaras E, Ylisaukko-oja T, Peltonen L, Jarvela I, Ioannidis JP: A heterogeneity-based genome search meta-analysis for autism-spectrum disorders. Mol Psychiatry 2006; 11:29–36
19. Xiong M, Guo SW: Fine-scale genetic mapping based on linkage disequilibrium: theory and applications. Am J Hum Genet 1997; 60:1513–1531
20. Norton N, Williams HJ, Owen MJ: An update on the genetics of schizophrenia. Curr Opin Psychiatry 2006; 19:158–164
21. Risch N, Merikangas K: The future of genetic studies of complex human diseases. Science 1996; 273:1516–1517
22. Pe'er I, Chretien YR, de Bakker PI, Barrett JC, Daly MJ, Altshuler DM: Biases and reconciliation in estimates of linkage disequilibrium in the human genome. Am J Hum Genet 2006; 78:588–603
23. Kruglyak L: The road to genome-wide association studies. Nat Rev Genet 2008; 9:314–318
24. Pritchard JK: Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 2001; 69:124–137
25. Guthery SL, Salisbury BA, Pungliya MS, Stephens JC, Bamshad M: The structure of common genetic variation in United States populations. Am J Hum Genet 2007; 81:1221–1231
26. Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, Lohmueller KE, Adams MD, Schmidt S, Sninsky JJ, Sunyaev SR, White TJ, Nielsen R, Clark AG, Bustamante CD: Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet 2008; 4:e1000083
27. Stone EA, Sidow A: Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. Genome Res 2005; 15:978–986
28. Duan J, Wainwright MS, Comeron JM, Saitou N, Sanders AR, Gelernter J, Gejman PV: Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. Hum Mol Genet 2003; 12:205–216
29. Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermüller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaöz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Löytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA; NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Shahab A, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JN, Yu Y, Ruan Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrímsdóttir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ: Identification and analysis of functional elements in 1% of the human genome by the ENCODE Pilot Project. Nature 2007; 447:799–816
30. Botstein D, Risch N: Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. Nat Genet 2003; 33(suppl):228–237
31. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH: Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 2004; 305:869–872
32. Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP: Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 2008; 40:592–599
33. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee A, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH: A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 2006; 314:1461–1463
34. Li M, Atmaca-Sonmez P, Othman M, Branham KE, Khanna R, Wade MS, Li Y, Liang L, Zareparsi S, Swaroop A, Abecasis GR: CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. Nat Genet 2006; 38:1049–1054
35. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, Inouye M, Freathy RM, Attwood AP, Beckmann JS, Berndt SI; Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, Jacobs KB, Chanock SJ, Hayes RB, Bergmann S, Bennett AJ, Bingham SA, Bochud M, Brown M, Cauchi S, Connell JM, Cooper C, Smith GD, Day I, Dina C, De S, Dermitzakis ET, Doney AS, Elliott KS, Elliott P, Evans DM, Sadaf Farooqi I, Froguel P, Ghori J, Groves CJ, Gwilliam R, Hadley D, Hall AS, Hattersley AT, Hebebrand J, Heid IM; KORA, Lamina C, Gieger C, Il-

lig T, Meitinger T, Wichmann HE, Herrera B, Hinney A, Hunt SE, Jarvelin MR, Johnson T, Jolley JD, Karpe F, Keniry A, Khaw KT, Luben RN, Mangino M, Marchini J, McArdle WL, McGinnis R, Meyre D, Munroe PB, Morris AD, Ness AR, Neville MJ, Nica AC, Ong KK, O'Rahilly S, Owen KR, Palmer CN, Papadakis K, Potter S, Pouta A, Qi L; Nurses' Health Study, Randall JC, Rayner NW, Ring SM, Sandhu MS, Scherag A, Sims MA, Song K, Soranzo N, Speliotes EK; Diabetes Genetics Initiative, Syddall HE, Teichmann SA, Timpson NJ, Tobias JH, Uda M; SardiNIA Study, Vogel CI, Wallace C, Waterworth DM, Weedon MN; Wellcome Trust Case Control Consortium, Willer CJ; FUSION, Wraight, Yuan X, Zeggini E, Hirschhorn JN, Strachan DP, Ouwehand WH, Caulfield MJ, Samani NJ, Frayling TM, Vollenweider P, Waeber G, Mooser V, Deloukas P, McCarthy MI, Wareham NJ, Barroso I, Jacobs KB, Chanock SJ, Hayes RB, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Kraft P, Hankinson SE, Hunter DJ, Hu FB, Lyon HN, Voight BF, Ridderstrale M, Groop L, Scheet P, Sanna S, Abecasis GR, Albai G, Nagaraja R, Schlessinger D, Jackson AU, Tuomilehto J, Collins FS, Boehnke M, Mohlke KL: Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet 2008; 40:768–775

36. Murphy KC, Jones LA, Owen MJ: High rates of schizophrenia in adults with velo-cardio-facial syndrome. Arch Gen Psychiatry 1999; 56:940–945
37. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME: Global variation in copy number in the human genome. Nature 2006; 444: 444–454
38. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006; 38: 904–909
39. Marchini J, Howie B, Myers S, McVean G, Donnelly P: A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007; 39:906–913
40. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN: Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008; 9:356–369
41. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661–678
42. Dudbridge F, Gusnanto A: Estimation of significance thresholds for genomewide association scans. Genet Epidemiol 2008; 32: 227–234
43. Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ: Genome-wide significance for dense SNP and resequencing data. Genet Epidemiol 2008; 32:179–185
44. Pe'er I, Yelensky R, Altshuler D, Daly MJ: Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 2008; 32:381–385
45. Sabatti C, Service S, Freimer N: False discovery rate in linkage and association genome screens for complex disorders. Genetics 2003; 164:829–833
46. Wakefield J: Bayes factors for genome-wide association studies: comparison with p-values. Genet Epidemiol 2008
47. Frayling TM: Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 2007; 8: 657–662
48. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D: Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008; 40:638–645
49. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007; 316:889–894
50. Freathy RM, Timpson NJ, Lawlor DA, Pouta A, Ben-Shlomo Y, Ruokonen A, Ebrahim S, Shields B, Zeggini E, Weedon MN, Lindgren CM, Lango H, Melzer D, Ferrucci L, Paolisso G, Neville MJ, Karpe F, Palmer CN, Morris AD, Elliott P, Jarvelin MR, Smith GD, McCarthy MI, Hattersley AT, Frayling TM: Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected, given its effect on BMI. Diabetes 2008; 57: 1419–1426
51. Shoulders CC: The FTO (fat mass and obesity-associated) gene: big in adipocyte lipolysis? J Lipid Res 2008; 49:495–496
52. Todd JA: Statistical false positive or true disease pathway? Nat Genet 2006; 38:731–733
53. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331–1336
54. Ghoussaini M, Song H, Koessler T, Al Olama AA, Kote-Jarai Z, Driver KE, Pooley KA, Ramus SJ, Kjaer SK, Hogdall E, DiCioccio RA, Whittemore AS, Gayther SA, Giles GG, Guy M, Edwards SM, Morrison J, Donovan JL, Hamdy FC, Dearnaley DP, Ardern-Jones AT, Hall AL, O'Brien LT, Gehr-Swain BN, Wilkinson RA, Brown PM, Hopper JL, Neal DE, Pharoah PD, Ponder BA, Eeles RA, Easton DF, Dunning AM: Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst 2008; 100:962–966

55. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigmundsson T, Buizer-Voskamp JE, Hansen T, Jakobsen KD, Muglia P, Francks C, Matthews PM, Gylfason A, Halldorsson BV, Gudbjartsson D, Thorgeirsson TE, Sigurdsson A, Jonasdottir A, Jonasdottir A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Giegling I, Moller HJ, Hartmann A, Shianna KV, Ge D, Need AC, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Paunio T, Toulopoulou T, Bramon E, Di Forti M, Murray R, Ruggeri M, Vassos E, Tosato S, Walshe M, Li T, Vasilescu C, Muhleisen TW, Wang AG, Ullum H, Djurovic S, Melle I, Olesen J, Kiemeney LA, Franke B, Kahn RS, Linszen D, van Os J, Wiersma D, Bruggeman R, Cahn W, Germeys I, de Haan L, Krabbendam L, Sabatti C, Freimer NB, Gulcher JR, Thorsteinsdottir U, Kong A, Andreassen OA, Ophoff RA, Georgi A, Rietschel M, Werge T, Petursson H, Goldstein DB, Nothen MM, Peltonen L, Collier DA, St Clair D, Stefansson K: Large recurrent microdeletions associated with schizophrenia. Nature 2008; 455:232–236

56. Stone JL, O'Donovan MC, Gurling H, Kirov GK, Blackwood DH, Corvin A, Craddock NJ, Gill M, Hultman CM, Lichtenstein P, McQuillin A, Pato CN, Ruderfer DM, Owen MJ, St Clair D, Sullivan PF, Sklar P, Purcell SM, Stone JL, Ruderfer DM, Korn J, Kirov GK, Macgregor S, McQuillin A, Morris DW, O'Dushlaine CT, Daly MJ, Visscher PM, Holmans PA, O'Donovan MC, Sullivan PF, Sklar P, Purcell SM, Gurling H, Corvin A, Blackwood DH, Craddock NJ, Gill M, Hultman CM, Kirov GK, Lichtenstein P, McQuillin A, O'Donovan MC, Owen MJ, Pato CN, Purcell SM, Scolnick EM, St Clair D, Stone JL, Sullivan PF, Sklar P, O'Donovan MC, Kirov GK, Craddock NJ, Holmans PA, Williams NM, Georgieva L, Nikolov I, Norton N, Williams H, Toncheva D, Milanova V, Owen MJ, Hultman CM, Lichtenstein P, Thelander EF, Sullivan P, Morris DW, O'Dushlaine CT, Kenny E, Waddington JL, Gill M, Corvin A, McQuillin A, Choudhury K, Datta S, Pimm J, Thirumalai S, Puri V, Krasucki R, Lawrence J, Quested D, Bass N, Curtis D, Gurling H, Crombie C, Fraser G, Kwan SL, Walker N, St Clair D, Blackwood DH, Muir WJ, McGhee KA, Pickard B, Malloy P, Maclean AW, Van Beck M, Visscher PM, Macgregor S, Pato MT, Medeiros H, Middleton F, Carvalho C, Morley C, Fanous A, Conti D, Knowles JA, Ferreira CP, Macedo A, Azevedo MH, Pato CN, Stone JL, Ruderfer DM, Korn J, McCarroll SA, Daly M, Purcell SM, Sklar P, Purcell SM, Stone JL, Chambert K, Ruderfer DM, Korn J, McCarroll SA, Gates C, Daly MJ, Scolnick EM, Sklar P: Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 2008; 455:178–179

57. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Roccanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen Z, Davis S, Baker C, Eichler EE, Meltzer PS, Nelson SF, Singleton AB, Lee MK, Rapoport JL, King MC, Sebat J: Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 2008; 320:539–543

58. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M: Strong association of de novo copy number mutations with sporadic schizophrenia. Nat Genet 2008; 40:880–885

59. Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, Nimgaonkar VL, McQueen MB, Faraone SV, Kirby A, de Bakker PI, Ogdie MN, Thase ME, Sachs GS, Todd-Brown K, Gabriel SB, Sougnez C, Gates C, Blumenstiel B, Defelice M, Ardlie KG, Franklin J, Muir WJ, McGhee KA, MacIntyre DJ, McLean A, VanBeck M, McQuillin A, Bass NJ, Robinson M, Lawrence J, Anjorin A, Curtis D, Scolnick EM, Daly MJ, Blackwood DH, Gurling HM, Purcell SM: Whole-genome association study of bipolar disorder. Mol Psychiatry 2008; 13:558–569

60. Ferreira Mar, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, Smoller JW, Grozeva D, Stone J, Nikolov I, Chambert K, Hamshere ML, Nimgaonkar V, Moskvina V, Thase ME, Caesar S, Sachs GS, Franklin J, Gordon-Smith K, Ardlie K, Gabriel SB, Fraser C, Blumenstiel B, Defelice M, Breen G, Gill M, Morris DW, Elkin A, Muir WJ, McGhee KA, Williamson R, MacIntyre D, McLean A, Clair D, VanBeck M, Pereira A, Kandaswamy R, McQuillin A, Collier DA, Bass NJ, Young AH, Lawrence J, Ferrier I, Anjorin A, Farmer A, Curtis D, Scolnick EM, McGuffin P, Daly MJ, Corvin AP, AHolmans P, Blackwood DH, Consortium WTCC, Gurling HM, Owen MJ, Purcell SM, Sklar P, Craddock N: Collaborative genome-wide association analysis of 10,596 individuals supports a role for Ankyrin-G (ANK3) and the alpha-1C subunit of the L-type voltage-gated calcium channel (CACNA1C) in bipolar disorder. Nat Genet 2008 (in press)

61. Lencz T, Morgan TV, Athanasiou M, Dain B, Reed CR, Kane JM, Kucherlapati R, Malhotra AK: Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. Mol Psychiatry 2007; 12:572–580

62. Sullivan PF, Lin D, Tzeng JY, van den Oord E, Perkins D, Stroup TS, Wagner M, Lee S, Wright FA, Zou F, Liu W, Downing AM, Lieberman J, Close SL: Genomewide association for schizophrenia in the CATIE study: results of stage 1. Mol Psychiatry 2008; 13: 570–584

63. O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, Nikolov I, Hamshere M, Carroll L, Georgieva L, Dwyer S, Holmans P, Marchini JL, Spencer CC, Howie B, Leung HT, Hartmann AM, Moller HJ, Morris DW, Shi Y, Feng G, Hoffmann P, Propping P, Vasilescu C, Maier W, Rietschel M, Zammit S, Schumacher J, Quinn EM, Schulze TG, Williams NM, Giegling I, Iwata N, Ikeda M, Darvasi A, Shifman S, He L, Duan J, Sanders AR, Levinson DF, Gejman PV, Gejman PV, Sanders AR, Duan J, Levinson DF, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Cloninger CR, Cichon S, Nothen MM, Gill M, Corvin A, Rujescu D, Kirov G, Owen MJ: Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat Genet 2008; 40:1053–1055

64. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimaki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M: Strong association of de novo copy number mutations with autism. Science 2007; 316:445–449

65. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH Jr, Dobyns WB, Christian SL: Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet 2008; 17:628–638

66. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW: Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 2008; 82: 477–488

67. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ: Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med 2008; 358:667–675

68. Christian SL, Brune CW, Sudi J, Kumar RA, Liu S, Karamohamed S, Badner JA, Matsui S, Conroy J, McQuaid D, Gergel J, Hatchwell E, Gilliam TC, Gershon ES, Nowak NJ, Dobyns WB, Cook EH Jr:

Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. Biol Psychiatry 2008; 63:1111–1117
69. Kirov G, Zaharieva I, Georgieva L, Moskvina V, Nikolov I, Cichon S, Hillmer A, Toncheva D, Owen MJ, O'Donovan MC: A genome-wide association study in 574 schizophrenia trios using DNA pooling. Mol Psychiatry 2008; Mar 11 [Epub ahead of print]
70. Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, Craddock NJ, Kendler KS, Li T, O'Donovan M, O'Neill FA, Owen MJ, Walsh D, Weinberger DR, Sun C, Flint J, Darvasi A: Genome-wide association identifies a common variant in the Reelin gene that increases the risk of schizophrenia only in women. PLoS Genet 2008; 4:e28
71. Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, Schulze TG, Cichon S, Rietschel M, Nothen MM, Georgi A, Schumacher J, Schwarz M, Abou Jamra R, Hofels S, Propping P, Satagopan J, Detera-Wadleigh SD, Hardy J, McMahon FJ: A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 2008; 13:197–207
72. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A: Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003; 60:837–844
73. Faraone SV, Sergeant J, Gillberg C, Biederman J: The worldwide prevalence of ADHD: is it an American condition? World Psychiatry 2003; 2:104–113
74. Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA, Sklar P: Molecular genetics of attention-deficit/hyperactivity disorder. Biol Psychiatry 2005; 57:1313–1323
75. Rutter M: Incidence of autism spectrum disorders: changes over time and their meaning. Acta Pediatr 2005; 94:2–15
76. Freitag CM: The genetics of autistic disorders and its clinical relevance: a review of the literature. Mol Psychiatry 2007; 12:2–22
77. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, Karam EG, Lee CK, Lellouch J, Lepine JP, Newman SC, Rubio-Stipec M, Wells JE, Wickramaratne PJ, Wittchen H, Yeh EK: Cross-national epidemiology of major depression and bipolar disorder. JAMA 1996; 276:293–299
78. McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A: The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. Arch Gen Psychiatry 2003; 60:497–502
79. Hasin DS, Goodwin RD, Stinson FS, Grant BF: Epidemiology of major depressive disorder: results from the national epidemiologic survey on alcoholism and related conditions. Arch Gen Psychiatry 2005; 62:1097–1106
80. Sullivan PF, Neale MC, Kendler KS: Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 2000; 157:1552–1562
81. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ: The lifetime history of major depression in women: reliability of diagnosis and heritability. Arch Gen Psychiatry 1993; 50:863–870
82. Thapar A, McGuffin P: A twin study of depressive symptoms in childhood. Br J Psychiatry 1994; 165:259–265
83. Saha S, Chant D, Welham J, McGrath J: A systematic review of the prevalence of schizophrenia. PLoS Med 2005; 2:e141
84. Sullivan PF, Kendler KS, Neale MC: Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry 2003; 60:1187–1192
85. McGuffin P, Katz R, Watkins S, Rutherford J: A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiatry 1996; 53:129–136
86. Achkar JP, Duerr R: The expanding universe of inflammatory bowel disease genetics. Curr Opin Gastroenterol 2008; 24:429–434
87. Gottesman II, Gould TD: The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003; 160:636–645
88. Marchini J, Donnelly P, Cardon LR: Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet 2005; 37:413–417
89. Ptak C, Petronis A: Epigenetics and complex disease: from etiology to new therapeutics. Annu Rev Pharmacol Toxicol 2008; 48:257–276
90. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559–575
91. Purcell S, Cherny SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003; 19:149–150